

Gestalt recognition in a virtual melody experiment

M. Schulte, A. Knief, A. Seither-Preisler, and C. Pantev

*Biomagnetism Center, Institute of Experimental Audiology
University of Münster, Germany*

1 Introduction

The pitch of a harmonic complex tone corresponds to its fundamental frequency (f_0). Interestingly, this sensation does not depend on the physical presence of f_0 . The phenomenon, commonly referred to as virtual pitch [1], has often been compared to the perception of visual illusory contours. It is contrasted by the perception of spectral pitches, elicited by a tone's physically present harmonics.

While, in principle, each human subject and even higher vertebrates [2,3,4] are able to perceive missing fundamentals, the predominance of one type of pitch perception over the other (spectral vs. virtual) has been shown to depend on the spectral content of the stimulus as well as on the subject's musical experience [5]. Experiments using ambiguous two-tone sequences revealed that "spectral listeners" can be trained to switch to the virtual mode of pitch perception. In contrast, for "virtual listeners", a reversed training was without success [6]. This finding indicates a hierarchical organization of the two processes, the spectral mode being the basis for higher level virtual pitch formation.

Thus, the aim of the present study was to investigate the role of the auditory cortex for both modes of pitch perception.

In the present experiment, auditory evoked fields (AEFs) in the low frequency range (0-20Hz, slowAEF, N1m component with latency of about 100ms after stimulus onset) and in the gamma frequency range (24-48Hz [8], with its maximum at latencies of about 30-80ms, c.f. figure 4) elicited by a sequence of complex tones with their fundamental frequencies missing were investigated. N1m and the gamma band responses were recorded from the left and the right hemisphere of the test subjects, when they:

- perceived the "foreground" spectral melody and
- when they recognized the "hidden" virtual melody after successful training

2 Methods

2.1 Experimental setup

A sequence of eight complex tones, which missing fundamentals encoded the beginning of the song 'Frere Jacques', was presented. The harmonics were selected such that the contour of the spectral melody was inverse to the sought tune (c.f. Fig. 1). Thus, the ambiguity of the melody was utilized as an indicator of virtual pitch perception or in other words, the auditory Gestalt recognition.

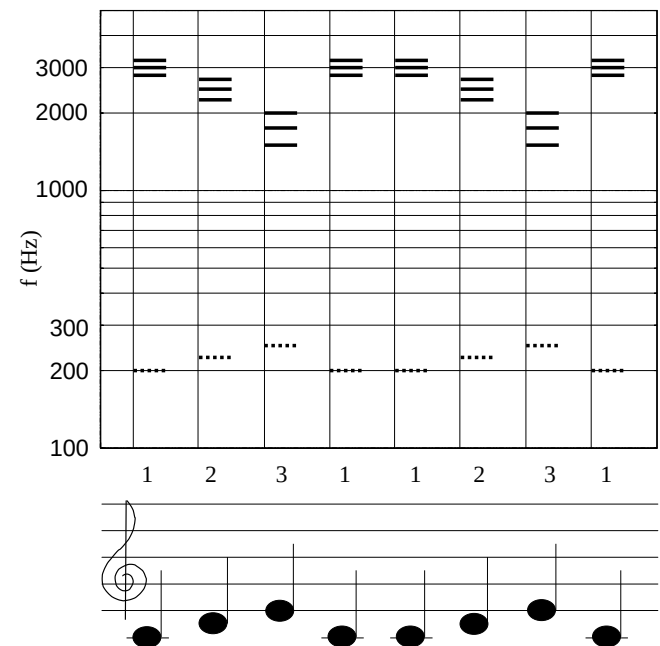


Figure 1: Stimulus sequence of the virtual melody experiment. The sequence was built up of three different complex tones ($f_{0,Tone1}$: 200 Hz, $f_{0,Tone2}$: 225 Hz and $f_{0,Tone3}$: 250 Hz), played in the order 12311231. The virtual melody follows the tune 'Frere Jacques' (lower inset: notation transposed to C major, upper inset: dotted lines), whereas the spectral melody, given by three upper harmonics, has an inverse pitch contour (upper inset: solid lines)

Eight subjects (five males and three females; right handed) who reported to perceive the spectral

melody, but not the virtual one, were measured by MEG (200 sequences per hemisphere). Thereafter, they had to listen to the melody for about one hour per day without any feedback until they reported a change in the perceived melody. In all subjects, due to training, pitch sensations switched to the virtual mode and the tune 'Frere Jacques' was clearly recognized. Subjects who did not succeed at the more difficult level (c.f. Table 1 below, first values of columns 2 and 3) after four hours, were trained at the less difficult level (c.f. Table1, values in brackets of columns 2 and 3). After successful training and perceiving the virtual melody, a second MEG measurement was performed in a passive listening experiment. During the MEG recording the subjects viewed a self selected video.

The three complex tones were adjusted for loudness (60dB sensation level). Stimulus duration and inter-stimulus interval were 150 ms and 650 ms, respectively.

Table 1: *Composition of the three complex tones. Each tone is composed of three harmonics (column 2) to a missing fundamental frequency (column 1). In brackets the harmonics for the less difficult level are shown.*

Missing fundamental [Hz]	Spectral content [Hz]	Harmonic rank
200	3200 (2800)	16 (14)
	3000 (2600)	15 (13)
	2800 (2400)	14 (12)
225	2700 (2475)	12 (11)
	2475 (2250)	11 (10)
	2250 (2025)	10 (9)
250	2000 (2000)	8 (8)
	1750 (1750)	7 (7)
	1500 (1500)	6 (6)

2.2 MEG measurements and source modelling

Magnetic field responses were recorded by a BTi 37 channel magnetometer. Source analysis, based on a moving Equivalent Current Dipole (ECD) model was applied to slow auditory evoked fields (sAEFs) and evoked gamma band fields (band-pass filter: 24-48 Hz).

For the sAEFs (N1m), stimulus related epochs to each of the three complex tones were averaged (tone 1: N = 800, tones 2 and 3: N = 400) and low-pass filtered at 20Hz. As a prerequisite for further analysis, a minimal goodness of fit (gof) of 90 % was required.

As can be seen in figure 2 and figure 4 gamma band AEFs had a lower field strength than sAEFs. Thus only average responses to tone 1 reached a sufficient gof of 85 %, because of a higher number of stimulus related epochs, and were considered for further analysis.

For statistical analysis, an ANOVA with training as a repeated measurements factor and hemisphere as an independent factor was performed.

3 Results

Four subjects were able to recognize the virtual melody in the more difficult task within four hours. The other four subjects recognized the melody at the more easy level within eight hours. No differences in results were found between these two groups.

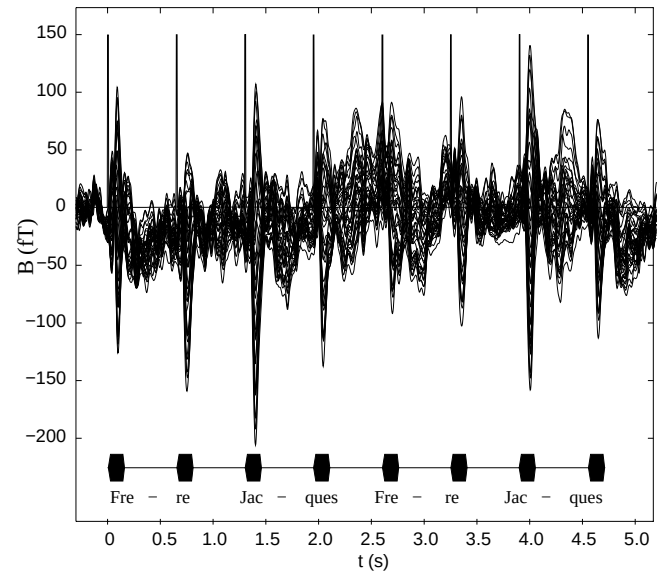


Figure 2: *Mean low-pass filtered responses to the eight tone-sequence of one individual subject (regular vertical lines: trigger signals).*

3.1 Slow auditory evoked fields (N1m)

Figure 2 illustrates the sAEF of one individual subject. Dipole modeling of the N1m yielded a sufficient gof of 90 % for 17 complex tone pairs in the left and for 18 complex tone pairs in the right hemisphere.

In this data (figure 3) no significant change in the dipole moment of the N1 equivalent source before and after the training was observed ($F=0.182$, $p=0.672$). The dipole moment decreased from 23.1 ± 2.5 nAm to 22.6 ± 3.9 nAm in the left and from 20.6 ± 2.3 nAm to 19.6 ± 1.6 nAm in the right hemisphere. Also the difference between hemispheres failed to reach significance ($F=0.655$, $p=0.424$).

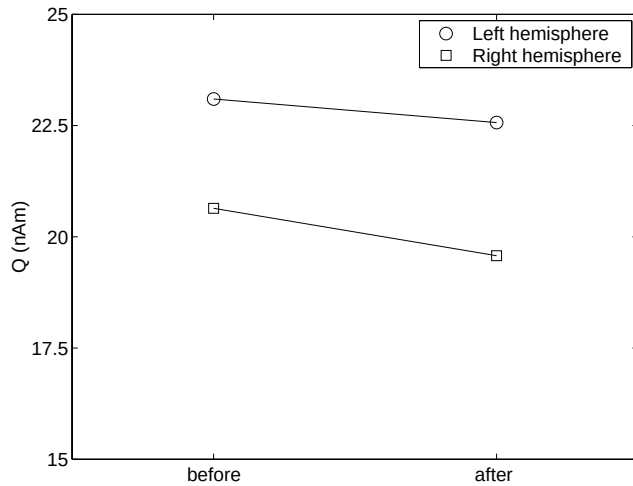


Figure 3: Dipole moment for N1m response before (spectral mode) and after training (virtual mode).

3.2 Transient gamma band responses

Gamma band responses obtained from four left and five right hemispheres in two corresponding sessions (before / after training) had a sufficient gof of at least 85 % (mean gof: 94%) and were considered for analysis.

Figure 4 shows data of an individual subject. Although the location of the sensor changed by only 0.3cm in radial direction, it is evident that after switching to the perception of virtual pitch due to the training, the gamma band response appreciatively increased.

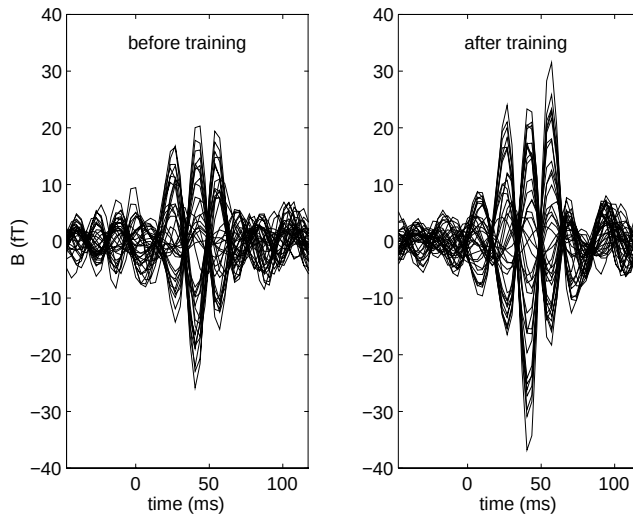


Figure 4: Example of band-pass filtered (24-48 Hz) gamma band responses, as obtained from one individual subject.

The switch from the spectral to the virtual type of pitch perception has also significantly increased the dipole moment of the gamma band source ($F=11.3$,

$p=0.012$) from 1.69 ± 0.23 nAm to 3.91 ± 0.79 nAm in the left and from 1.46 ± 0.18 nAm to 2.29 ± 0.53 nAm in the right hemisphere.

However, differences between the two hemispheres did not reach significance ($F=3.3$, $p=0.11$).

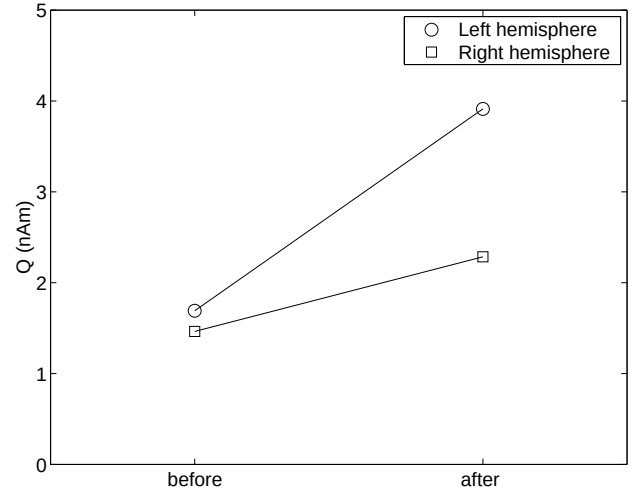


Figure 5: Dipole moment for gamma band responses before and after training,

4 Discussion

In all eight subjects, training resulted in a sudden switch from the spectral to the virtual mode of pitch perception.

The recognition of the virtual melody occurred abruptly and - analogous to the visual Gestalt recognition process - was irreversible. This finding indicates that

- virtual pitch sensations may reflect a higher order abstraction process, based on the more elementary mode of spectral pitch perception and
- plastic neural processes may be involved in virtual pitch formation.

In this data the recognition of the virtual melody was not associated with a change in the strength of N1m cortical source. This relative stability of the N1m dipole moment indicates, however, that an unspecific repetition effect cannot account for the significant increase of the evoked gamma band response ($p=0.012$). Rather this effect may be attributed to specific learning in recognizing the virtual pitch after some training.

The data suggest the possibility, that integration of different spectral pitches into a coherent virtual pitch percept may be associated with an increase of

synchronization or/and to an enlargement of the cortical network generating the evoked gamma band response.

The present findings are in agreement with those of an earlier study [8], where the gamma band has been found to be more synchronous for sinusoids and harmonic complex tones than for inharmonic tones. Therefore, we consider the gamma band to be a direct correlate of virtual pitch formation. Since the latency of the maximum of the transient gamma band oscillation is between 30-70ms, this process is likely to take place before or at the level of the primary auditory cortex.

Acknowledgements

The work was supported by the Deutsche Forschungsgemeinschaft Pa/392/4-3.

References

1. E. Terhardt, "Pitch, consonance, and harmony", *J. Acoust. Soc. Am.*, Vol. 55(5), 1974, pp. 1061-1069.
2. J. Cynx, M. Shapiro, "Perception of the missing fundamental by a species of songbird, the European starling (*Sturnus vulgaris*)", *Journal of Comparative Psychology*, Vol. 100, 1986, pp. 356-360.
3. A. Preisler, S. Schmidt, "Spontaneous classification of complex tones at high and ultrasonic frequencies in the bat, *Megaderma lyra*", *J. Acoust. Soc. Am.*, Vol. 103, 1998, pp. 2595-2607.
4. W. Tomlinson, D. Schwartz, "Perception of the missing fundamental in nonhuman primates", *J. Acoust. Soc. Am.*, Vol. 84, 1988, pp. 560-565.
5. A. Preisler, "The influence of spectral composition of complex tones and of musical experience on the perceptibility of virtual pitch", *Perception and Psychophysics*, Vol. 54(5), 1992, pp. 589-603.
6. W. Deutsch, R. Steeger, "Binaurales Residuum versus Lokalisation", *Mitteilungen der Kommission für Schallforschung, Österreichische Akademie der Wissenschaften*, Nr. 5, 1978, pp. 41-53.
7. C. Pantev, S. Makeig, M. Hoke, R. Galambos, S. Hampson, C. Gallen, "Human auditory evoked gamma-band magnetic fields", *Proc. Natl. Acad. Sci USA*, Vol. 88, 1991, pp. 8996-9000.
8. A. Knief, M. Schulte, O. Bertrand, and C. Pantev, "The perception of coherent and non-coherent auditory objects: a signature in gamma frequency band", *Hearing Research*, Vol. 145, 2000, pp. 161-168.